



# Improving NASA Earth science data and information access through natural language processing based data analysis and visualization

Zhong Liu<sup>1,2</sup>, Shimei Pan<sup>3</sup>, Jianwu Wang<sup>3</sup>, Angela Li<sup>1</sup>, and David Meyer<sup>1</sup>

<sup>1</sup>NASA Goddard Earth Sciences (GES) Data and Information Services Center (DISC)

<sup>2</sup>CSISS, George Mason University

<sup>3</sup>Dept. of Information Systems, UMBC



# Motivation

## NASA:

- The Research Access initiative is part of the agency's framework for increasing public access to scientific publications and digital scientific data.
- The initiative follows the release of White House Office of Science and Technology Policy's (OSTP) memorandum "[Increasing Access to the Results of Federally Funded Research](#)," to ensure federally funded research is available to the public within one year of publication.
- NASA answered the mandate by creating an agency plan entitled "[NASA Plan for Increasing Access to the Results of Scientific Research](#)" and associated policy, [NPD 2230.1, Research Data and Publication Access](#).

## Principles in NASA SMD Strategic Plan for Scientific Data and Computing:

- Continued free and open access to scientific data for any use
- Improved ease of use and discoverability
- Enhanced science applications and new use cases
- Incorporates best practices and "state of the art" through partnerships

## Earth Data and Systems are Evolving:

- Increasing archive and file sizes. More complicated data structures
- More user-friendly and data services
- What is the future direction?



# Motivation (cont.)

## Challenges in data access:

- “Decision Support Systems Analysts, the General Public, and University Undergraduates report the lowest levels of CSI” according to the 2017 CSI (Customer Survey Index). Over 50% of users.
- Surveys reveal that most non-professional users normally do not want to download and handle raw data as well as conduct heavy-duty data processing tasks.

	2016	2016	2016	2017	2017	2017
	%	N	CSI	%	N	CSI
Type of User~						
General Public	14%	1,019	76	14%	1,037	76
Elementary, Middle, High School Teacher	1%	83	76	1%	86	77
University Professor	16%	1,129	80	16%	1,193	81
University Undergraduate Student	36%	2,550	76	9%	656	76
Other Education and Outreach	5%	349	79	5%	355	79
Earth Science Researcher	32%	2,304	79	32%	2,409	79
Earth Science Modeler	8%	574	78	9%	650	79
NASA-affiliated Scientist	2%	167	79	1%	102	80
Non-NASA-affiliated Scientist	4%	304	79	4%	320	78
NASA Science Team Member	7%	475	79	1%	68	80
Data Tool Developer/Provider	5%	359	77	5%	409	77
Decision Support Systems Analyst	5%	375	76	6%	429	76
University Graduate Student	0%	0	--	29%	2,204	77
Other User Type	8%	548	76	9%	656	77
<b>Number of Respondents</b>	<b>7,133</b>	<b>7,133</b>	<b>7,133</b>	<b>7,505</b>	<b>7,505</b>	<b>7,505</b>

*2017 ACSI  
Survey Results*



# Motivation (cont.)

*Kindly allow 10-15 business days for processing*

Ask NASA:

Please fill out this form completely, then click "Submit."

So that we can sort e-mail and respond more effectively, please include a specific subject line. Messages with blank or general subject lines such as "I have a question" or "Hello" will be read last. Kindly allow 10 to 15 business days for processing.

First Name:

Last Name:

E-mail Address:

Subject:

Ask NASA:

Please enter your question or comment here.

Submit

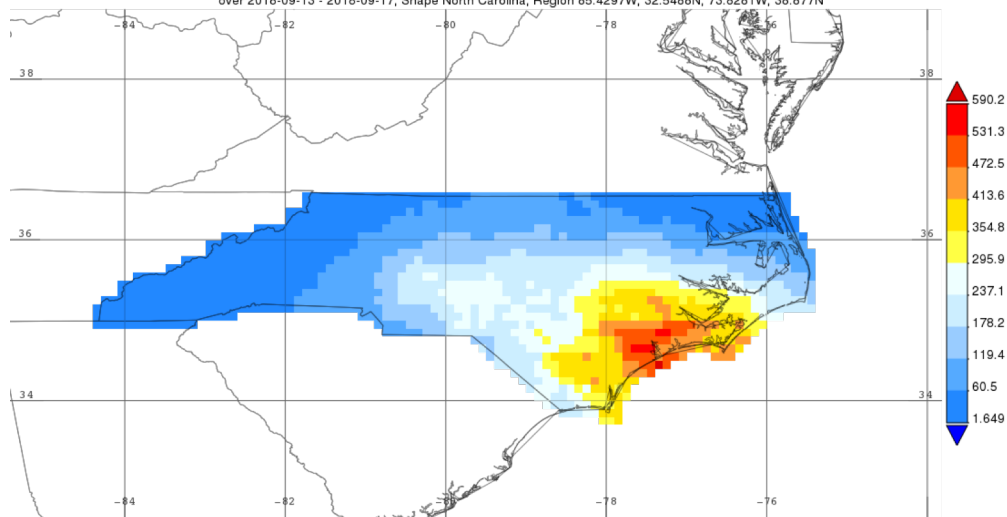


**Ask NASA**

Map of total rainfall from Hurricane Florence in North Carolina?



Map, Accumulated of Daily accumulated precipitation (combined microwave-IR) estimate - Late Run daily 0.1 deg. [GPM GPM\_3IMERGDL v05] mm over 2018-09-13 - 2018-09-17, Shape North Carolina, Region 85.4297W, 32.5488N, 73.8281W, 38.877N



[https://www.nasa.gov/about/contact/ask\\_nasa\\_form.html](https://www.nasa.gov/about/contact/ask_nasa_form.html)

Google

map of total rainfall from Hurricane Florence in North Carolina?

All Maps Images News Shopping More Settings Tools

About 748,000 results (0.58 seconds)

**Hurricane Florence: rainfall totals for coastal areas across NC | Charlotte ...**  
<https://www.charlotteobserver.com/news/states/north-carolina/article218495840.html>  
Sep 16, 2018 - Florence has brought more than 10 inches of rain on parts of Eastern ... Florence rain gauge map: Totals for more than 10 inches of rain on parts of Eastern ...

**Hurricane Florence: rainfall totals for coastal areas across NC, SC tropical storm | News ...**  
<https://www.newsobserver.com/news/states/north-carolina/article218446410.html>  
Sep 15, 2018 - US Geological Survey stations show record amounts of rainfall since Hurricane Florence reached the coast. ... continued to ...

**Florence Sets Preliminary North Carolina and South Carolina Tropical ...**  
<https://weather.com/news/stories/north-carolina-tropical-rain-2018-09-15>  
Sep 15, 2018 - Florence sets Tropical Rainfall Record for North Carolina and South Carolina ... Rainfall totals from Hurricane Florence from Sep. 13-17, 2018.

**Florence tracker: latest maps show hurricane path and rainfall | World ...**  
<https://www.theguardian.com/us-news/2018/sep/17/florence-hurricane-path-where-is-storm-surge-headin->  
Sep 17, 2018 - Hurricane Florence made landfall near Wilmington, North Carolina on Friday ... amounts of rainfall will continue to flood large parts of North ...



# Solution and Activities

## Solution:

- Develop natural language processing (NLP) based data analysis and visualization infrastructure

Input (text, voice)



Processing (text, voice)



Analysis and Visualization



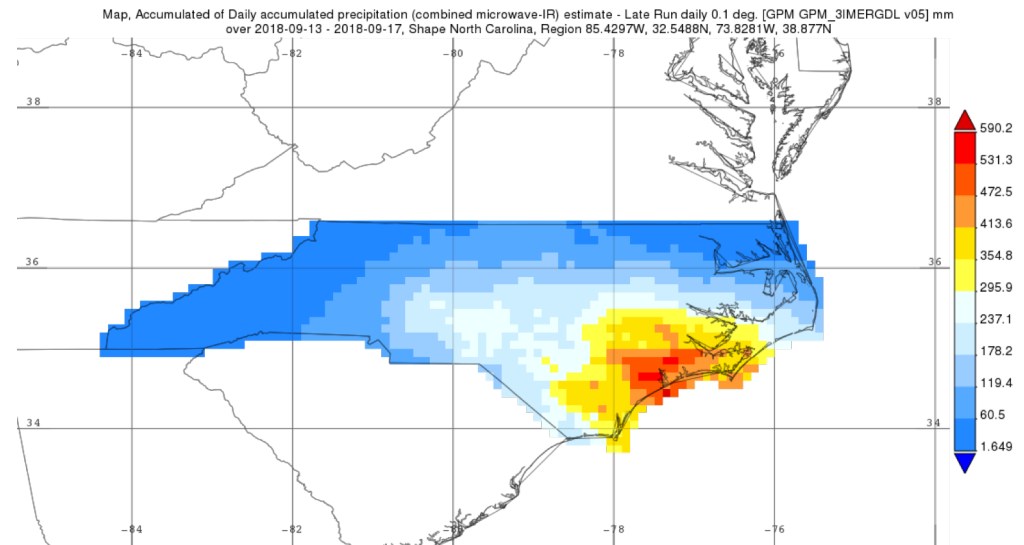
Result (text, voice, graphics)

Ask **NASA**

What is the total rainfall from Hurricane Florence in North Carolina?



- Collect user input info. about where, when, what, etc. (total rainfall map, Hurricane Florence, North Carolina)
- Call a backend system to process the inputs and generate the result (the rainfall map)



## Activities:

- Work with NLP experts at UMBC
- Use case development
- System design and prototyping



# Summary

- Many challenges in Earth science data and information access for users at all levels
- NLP provides a simple (but difficult to develop) interface to ordinary users
- NLP provides inputs for backend processing (data analysis and visualization)
- Working with NLP experts to develop a prototype